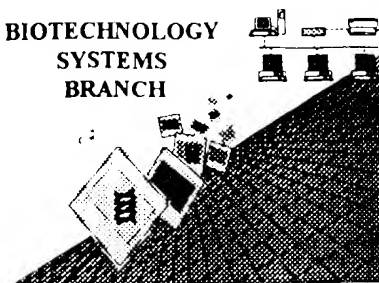


## **RAW SEQUENCE LISTING** **ERROR REPORT**

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



CP

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/743,247  
Source: Per/09  
Date Processed by STIC: 8/3/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST 25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

# Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/743,247
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 _____ Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 _____ Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 _____ Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 _____ Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 _____ Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 _____ PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 _____ Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 _____ Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 _____ Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 _____ Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 _____ Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 _____ PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 _____ Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.	

PCT09

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,247

DATE: 08/03/2001

TIME: 16:29:03

Input Set : A:\ES.txt

Output Set: N:\CRF3\08032001\I743247.raw

3 <110> APPLICANT: Sagami Chemical Research Center; Protegene Inc.  
 5 <120> TITLE OF INVENTION: Human Proteins Having Hydrophobic Domains And DNAs Encoding  
 These  
 6 Proteins  
 8 <130> FILE REFERENCE: 661102  
 10 <140> CURRENT APPLICATION NUMBER: US/09/743,247  
 10 <141> CURRENT FILING DATE: 1999-07-22  
 10 <150> PRIOR APPLICATION NUMBER: JP 10-208820  
 11 <151> PRIOR FILING DATE: 1998-07-24  
 13 <150> PRIOR APPLICATION NUMBER: JP 10-224105  
 14 <151> PRIOR FILING DATE: 1998-08-07  
 16 <150> PRIOR APPLICATION NUMBER: JP 10-238116  
 17 <151> PRIOR FILING DATE: 1998-08-25  
 19 <150> PRIOR APPLICATION NUMBER: JP 10-254736  
 20 <151> PRIOR FILING DATE: 1998-09-09  
 22 <150> PRIOR APPLICATION NUMBER: JP 10-275505  
 23 <151> PRIOR FILING DATE: 1998-09-29  
 25 <160> NUMBER OF SEQ ID NOS: 150  
 27 <170> SOFTWARE: Windows 95 (Word 98)

## ERRORED SEQUENCES

519 <210> SEQ ID NO: 21  
 520 <211> LENGTH: (510) 509 (p 2)  
 521 <212> TYPE: DNA  
 522 <213> ORGANISM: Homo sapiens  
 523 <220> FEATURE:  
 524 <221> NAME/KEY: CDS  
 525 <222> LOCATION: (66)...(443)  
 527 <400> SEQUENCE: 21  
 528 acgcttgatc cccggccgcg gggccaggaa gtcggagttt gagccccgga ggcagagcgg 60  
 529 ctgcc atg gcc aag tac ctg gcc cag atc att gtg atg gcc gtg cag gtg 110  
 530 Met Ala Lys Tyr Leu Ala Gln Ile Ile Val Met Gly Val Gln Val  
 531 1 5 10 15  
 532 gtg gcc agg gcc ttt gca cgg gcc ttg cgg cag gag ttt gca gcc agc 158  
 533 Val Gly Arg Ala Phe Ala Arg Ala Leu Arg Gln Glu Phe Ala Ala Ser  
 534 20 25 30  
 535 cgg gcc gca gct gat gcc cga gga cgc gct gga cac cgg tct gca gcc 206  
 536 Arg Ala Ala Ala Asp Ala Arg Gly Arg Ala Gly His Arg Ser Ala Ala  
 537 35 40 45  
 538 gct tcc aac ctc tcc ggc ctc agc ctc cag gag gca cag cag att ctc 254  
 539 Ala Ser Asn Leu Ser Gly Leu Ser Leu Gln Glu Ala Gln Gln Ile Leu  
 540 50 55 60  
 541 aac gtg tcc aag ctg agc cct gag gag gtc cag aag aac tat gaa cac 302  
 542 Asn Val Ser Lys Leu Ser Pro Glu Glu Val Gln Lys Asn Tyr Glu His  
 543 65 70 75  
 544 tta ttt aag gtg aat gat aaa tcc gtg ggt ggc tcc ttc tac ctg cag 350

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,247

DATE: 08/03/2001

TIME: 16:29:03

Input Set : A:\ES.txt

Output Set: N:\CRF3\08032001\I743247.raw

```

545 Leu Phe Lys Val Asn Asp Lys Ser Val Gly Gly Ser Phe Tyr Leu Gln
546 80 85 90 95
547 tca aag gtg gtc cgc gca aag gag cgc ctg gat gag gaa ctc aaa atc
548 Ser Lys Val Val Arg Ala Lys Glu Arg Leu Asp Glu Glu Leu Lys Ile
549 100 105 110
E--> 550 cag gcc cag gag gac aga gaa aaa ggg cag atg ccc cat acg tgactgctc
551 Gln Ala Gln Glu Asp Arg Glu Lys Gly Gln Met Pro His Thr
552 115 120 125
E--> 553 gctccccccg cccaccccg cgcctctaatt ttatagcttg gtaataaatt tcttttctgc
3553 <210> SEQ ID NO: 93
3554 <211> LENGTH: 195
3555 <212> TYPE: PRT
3556 <213> ORGANISM: Homo (sapiens)
3558 <400> SEQUENCE: 93
3559 Met Arg Leu Leu Leu Leu Leu Val Ala Ala Ser Ala Met Val Arg
3560 1 5 10 15
3561 Ser Glu Ala Ser Ala Asn Leu Gly Gly Val Pro Ser Lys Arg Leu Lys
3562 20 25 30
3563 Met Gln Tyr Ala Thr Gly Pro Leu Leu Lys Phe Gln Ile Cys Val Ser
3564 35 40 45
E--> 3565 Xaa Gly Tyr Arg Arg Val Phe Glu Glu Tyr Met Arg Val Ile Ser Gln
3566 50 55 60
3567 Arg Tyr Pro Asp Ile Arg Ile Glu Gly Glu Asn Tyr Leu Pro Gln Pro
3568 65 70 75 80
3569 Ile Tyr Arg His Ile Ala Ser Phe Leu Ser Val Phe Lys Leu Val Leu
3570 85 90 95
3571 Ile Gly Leu Ile Ile Val Gly Lys Asp Pro Phe Ala Phe Phe Gly Met
3572 100 105 110
3573 Gln Ala Pro Ser Ile Trp Gln Trp Gly Gln Glu Asn Lys Val Tyr Ala
3574 115 120 125
3575 Cys Met Met Val Phe Phe Leu Ser Asn Met Ile Glu Asn Gln Cys Met
3576 130 135 140
3577 Ser Thr Gly Ala Phe Glu Ile Thr Leu Asn Asp Val Pro Val Trp Ser
3578 145 150 155 160
3579 Lys Leu Glu Ser Gly His Leu Pro Ser Met Gln Gln Leu Val Gln Ile
3580 165 170 175
3581 Leu Asp Asn Glu Met Lys Leu Asn Val His Met Asp Ser Ile Pro His
3582 180 185 190
3583 His Arg Ser
3584 195
4525 <210> SEQ ID NO: 116
4526 <211> LENGTH: 1357
4527 <212> TYPE: DNA
4528 <213> ORGANISM: Homo (Sapiens)
WtE> 4529 <220> FEATURE:
4530 <221> NAME/KEY: CDS
4531 <222> LOCATION: (81)...(1262)
4533 <400> SEQUENCE: 116
4534 cgtgcgtttg tggccgtccg gcttcctga catgcagccc tctggacccc gaggttggac
60

```

398

(450) H49

(510) S09

→ see  
item 9  
on Enu  
summary  
sheet

Sapiens

P S

## RAW SEQUENCE LISTING

DATE: 08/03/2001

PATENT APPLICATION: US/09/743,247

TIME: 16:29:04

Input Set : A:\ES.txt

Output Set: N:\CRF3\08032001\I743247.raw

4535 cctactgtga cacacctacc atg cgg aca ctc ttc aac ctc ctc tgg ctt 110  
 4536 Met Arg Thr Leu Phe Asn Leu Leu Trp Leu  
 4537 1 5 10  
 4538 gcc ctg gcc tgc agc cct gtt cac act acc ctg tca aag tca gat gcc 158  
 4539 Ala Leu Ala Cys Ser Pro Val His Thr Thr Leu Ser Lys Ser Asp Ala  
 4540 15 20 25  
 E--> 4541 aaa aaa gcc gcc tca aag acg ctg ctg gag aag agt cag ttt tca gat  
 4542 Lys Lys Ala Ala Ser Lys Thr Leu Leu Glu Lys Ser Gln Phe Ser Asp  
 4543 30 35 40  
 4544 aag ccg gtg caa gac cgg ggt ttg gtg gtg acg gac ctc aaa gct gag 254  
 4545 Lys Pro Val Gln Asp Arg Gly Leu Val Val Thr Asp Leu Lys Ala Glu  
 4546 45 50 55  
 4547 agt gtg gtt ctt gag cat cgc agc tac tgc tgc gca aag gcc cgg gac 302  
 4548 Ser Val Val Leu Glu His Arg Ser Tyr Cys Ser Ala Lys Ala Arg Asp  
 4549 60 65 70  
 4550 aga cac ttt gct ggg gat gta ctg ggc tat gtc act cca tgg aac agc 350  
 4551 Arg His Phe Ala Gly Asp Val Leu Gly Tyr Val Thr Pro Trp Asn Ser  
 4552 75 80 85 90  
 4553 cat ggc tac gat gtc acc aag gtc ttt ggg agc aag ttc aca cag atc 398  
 4554 His Gly Tyr Asp Val Thr Lys Val Phe Gly Ser Lys Phe Thr Gln Ile  
 4555 95 100 105  
 4556 tca ccc gtc tgg ctg cag ctg aag aga cgt ggc cgt gag atg ttt gag 446  
 4557 Ser Pro Val Trp Leu Gln Leu Lys Arg Arg Gly Arg Glu Met Phe Glu  
 4558 110 115 120  
 4559 gtc acg ggc ctc cac gac gtg gac caa ggg tgg atg cga gct gtc agg 494  
 4560 Val Thr Gly Leu His Asp Val Asp Gln Gly Trp Met Arg Ala Val Arg  
 4561 125 130 135  
 4562 aag cat gcc aag ggc ctg cac ata gtg cct cgg ctc ctg ttt gag gac 542  
 4563 Lys His Ala Lys Gly Leu His Ile Val Pro Arg Leu Leu Phe Glu Asp  
 4564 140 145 150  
 4565 tgg act tac gat gat ttc cgg aac gtc tta gac agt gag gat gag ata 590  
 4566 Trp Thr Tyr Asp Asp Phe Arg Asn Val Leu Asp Ser Glu Asp Glu Ile  
 4567 155 160 165 170  
 4568 gag gag ctg agc aag acc gtg gtc cag gtg gca aag aac cag cat ttc 638  
 4569 Glu Glu Leu Ser Lys Thr Val Val Gln Val Ala Lys Asn Gln His Phe  
 4570 175 180 185  
 4571 gat ggc ttc gtg gtg gag gtc tgg aac cag ctg cta agc cag aag cgc 686  
 4572 Asp Gly Phe Val Val Glu Val Trp Asn Gln Leu Leu Ser Gln Lys Arg  
 4573 190 195 200  
 4574 gtg ggc ctc atc cac atg ctc acc cac ttg gcc gag gct ctg cac cag 734  
 4575 Val Gly Leu Ile His Met Leu Thr His Leu Ala Glu Ala Leu His Gln  
 4576 205 210 215  
 4577 gcc cgg ctg ctg gcc ctc ctg gtc atc ccg cct gcc atc acc ccc ggg 782  
 4578 Ala Arg Leu Leu Ala Leu Leu Val Ile Pro Pro Ala Ile Thr Pro Gly  
 4579 220 225 230  
 4580 acc gac cag ctg ggc atg ttc acg cac aag gag ttt gag cag ctg gcc 830  
 4581 Thr Asp Gln Leu Gly Met Phe Thr His Lys Glu Phe Glu Gln Leu Ala  
 4582 235 240 245 250  
 4583 ccc gtg ctg gat ggt ttc agc ctc atg acc tac gac tac tct aca gcg 878

(203) 206

## RAW SEQUENCE LISTING

DATE: 08/03/2001

PATENT APPLICATION: US/09/743,247

TIME: 16:29:04

Input Set : A:\ES.txt

Output Set: N:\CRF3\08032001\I743247.raw

```

4584 Pro Val Leu Asp Gly Phe Ser Leu Met Thr Tyr Asp Tyr Ser Thr Ala
4585                               255                               260                               265
4586 cat cag cct ggc cct aat gca ccc ctg tcc tgg gtt cga gcc tgc gtc      926
4587 His Gln Pro Gly Pro Asn Ala Pro Leu Ser Trp Val Arg Ala Cys Val
4588                               270                               275                               280
4589 cag gtc ctg gac ccg aag tcc aag tgg cga agc aaa atc ctc ctg ggg      974
4590 Gln Val Leu Asp Pro Lys Ser Lys Trp Arg Ser Lys Ile Leu Leu Gly
4591                               285                               290                               295
4592 ctc aac ttc tat ggt atg gac tac gcg acc tcc aag gat gcc cgt gag      1022
4593 Leu Asn Phe Tyr Gly Met Asp Tyr Ala Thr Ser Lys Asp Ala Arg Glu
4594                               300                               305                               310
4595 cct gtt gtc ggg gcc agg tac atc cag aca ctg aag gac cac agg ccc      1070
4596 Pro Val Val Gly Ala Arg Tyr Ile Gln Thr Leu Lys Asp His Arg Pro
4597 315                               320                               325                               330
4598 cgg atg gtg tgg gac agc cag gcc tca gag cac ttc ttc gag tac aag      1118
4599 Arg Met Val Trp Asp Ser Gln Ala Ser Glu His Phe Phe Glu Tyr Lys
4600                               335                               340                               345
4601 aag agc cgc agt ggg agg cac gtc gtc ttc tac cca acc ctg aag tcc      1166
4602 Lys Ser Arg Ser Gly Arg His Val Val Phe Tyr Pro Thr Leu Lys Ser
4603                               350                               355                               360
4604 ctg cag gtg cgg ctg gag ctg gcc cgg gag ctg ggc gtt ggg gtc tct      1214
4605 Leu Gln Val Arg Leu Glu Leu Ala Arg Glu Leu Gly Val Gly Val Ser
4606                               365                               370                               375
4607 atc tgg gag ctg ggc cag ggc ctg gac tac ttc tac gac ctg ctc t      1260
4608 Ile Trp Glu Leu Gly Gln Gly Leu Asp Tyr Phe Tyr Asp Leu Leu
4609                               380                               385                               390
4610 aggtgggcat tgcggcctcc gcggtggacg tggtcttttc taagccatgg agtgagtgag      1320
4611 caggtgtgaa atacaggcct ccactccgtt tgctgtg      1357
5735 <210> SEQ ID NO: 144
5736 <211> LENGTH: 619
5737 <212> TYPE: DNA
5738 <213> ORGANISM: Homo (sapience)
(WK-) 5739 <220> FEATURE:
5740 <221> NAME/KEY: CDS
5741 <222> LOCATION: (13)...(333)
5743 <400> SEQUENCE: 144
5744 cttcgactcg ct atg tcc act aac aat atg tgc gac cca cgg agg ccg      48
5745                               Met Ser Thr Asn Asn Met Ser Asp Pro Arg Arg Pro
5746                               1                               5                               10
5747 aac aaa gtg ctg agg tac aag ccc ccg ccg agc gaa tgt aac ccg gcc      96
5748 Asn Lys Val Leu Arg Tyr Lys Pro Pro Pro Ser Glu Cys Asn Pro Ala
5749                               15                               20                               25
5750 ttg gac gac ccg acg ccg gac tac atg aac ctg ctg ggc atg atc ttc      144
5751 Leu Asp Asp Pro Thr Pro Asp Tyr Met Asn Leu Leu Gly Met Ile Phe
5752                               30                               35                               40
5753 agc atg tgc ggc ctc atg ctt aag ctg aag tgg tgt gct tgg gtc gct      192
5754 Ser Met Cys Gly Leu Met Leu Lys Leu Lys Trp Cys Ala Trp Val Ala
5755 45                               50                               55                               60
5756 gtc tac tgc tcc ttc atc agc ttt gcc aac tct cgg agc tgc gag gac      240

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,247

DATE: 08/03/2001

TIME: 16:29:04

Input Set : A:\ES.txt

Output Set: N:\CRF3\08032001\I743247.raw

```

5757 Val Tyr Cys Ser Phe Ile Ser Phe Ala Asn Ser Arg Ser Ser Glu Asp
5758                65                70                75
5759 acg aag caa atg atg agt agc ttc atg ctg tcc atc tct gcc gtg gtg      288
5760 Thr Lys Gln Met Met Ser Ser Phe Met Leu Ser Ile Ser Ala Val Val
5761                80                85                90
E--> 5762 atg tcc tat ctg cag aat cct cag ccc atg acg ccc cca tgg      (340) 330
5763 Met Ser Tyr Leu Gln Asn Pro Gln Pro Met Thr Pro Pro Trp
5764                95                100                105
5765 tgataccagc ctagaagggt cacatttttg accctgtcta tccactaggc ctgggctttg      390
5766 gctgctaaac ctgctgccct cagctgccat cctggacttc cctgaatgag gccgtctcgg      450
5767 tgccccagc tggatagagg gaacctggcc ctttccctagg gaacacccta ggcttacccc      510
5768 tctgacctcc cttccctcgc ctgctgctgg gggagatgct gtccatgttt ctaggggtat      570
5769 tcatttgctt tctcgttgaa acctgttggt aataaagttt ttcactcag      619
5976 <210> SEQ ID NO: 149
5977 <211> LENGTH: 2182
5978 <212> TYPE: DNA
5979 <213> ORGANISM: Homo (sapience)
WtE> 5980 <220> FEATURE:
5981 <221> NAME/KEY: CDS
5982 <222> LOCATION: (56)...(1090)
5984 <400> SEQUENCE: 149
E--> 5985 gcacttcagc ttccctccc ccggcgccct ctggggctcc gagcccgcg ggacc      (58) 55
5986 atg ttc acc agc acc ggc tcc agt ggg ctc tac aag gcg cct ctg tcg      103
5987 Met Phe Thr Ser Thr Gly Ser Ser Gly Leu Tyr Lys Ala Pro Leu Ser
5988 1                5                10                15
5989 aag agc ctt ctg ctg gtc ccc agt gcc ctc tcc ctc ctg ctc gcc ctc      151
5990 Lys Ser Leu Leu Leu Val Pro Ser Ala Leu Ser Leu Leu Leu Ala Leu
5991                20                25                30
5992 ctc ctg cct cac tgc cag aag ctc ttt gtg tat gac ctt cac gca gtc      199
5993 Leu Leu Pro His Cys Gln Lys Leu Phe Val Tyr Asp Leu His Ala Val
5994                35                40                45
5995 aag aac gac ttc cag att tgg agg ttg ata tgt gga aga ata att tgc      247
5996 Lys Asn Asp Phe Gln Ile Trp Arg Leu Ile Cys Gly Arg Ile Ile Cys
5997 50                55                60
5998 ctt gat ttg aaa gat act ttc tgc agt agt ctg ctt att tat aat ttt      295
5999 Leu Asp Leu Lys Asp Thr Phe Cys Ser Ser Leu Leu Ile Tyr Asn Phe
6000 65                70                75                80
6001 agg ata ttt gaa aga aga tat gga agc aga aaa ttt gca tcc ttt ttg      343
6002 Arg Ile Phe Glu Arg Arg Tyr Gly Ser Arg Lys Phe Ala Ser Phe Leu
6003                85                90                95
6004 ctg ggt tcc tgg gtt ttg tca gcc tta ttt gac ttt ctc ctc att gaa      391
6005 Leu Gly Ser Trp Val Leu Ser Ala Leu Phe Asp Phe Leu Leu Ile Glu
6006                100                105                110
6007 gct atg cag tat ttc ttt ggc atc act gca gct agt aat ttg cct tct      439
6008 Ala Met Gln Tyr Phe Phe Gly Ile Thr Ala Ala Ser Asn Leu Pro Ser
6009                115                120                125
6010 gga ttc ctg gca cct gtg ttt gct ctg ttt gta cca ttt tac tgc tcc      487
6011 Gly Phe Leu Ala Pro Val Phe Ala Leu Phe Val Pro Phe Tyr Cys Ser
6012                130                135                140

```

## RAW SEQUENCE LISTING

DATE: 08/03/2001

PATENT APPLICATION: US/09/743,247

TIME: 16:29:05

Input Set : A:\ES.txt

Output Set: N:\CRF3\08032001\I743247.raw

6013	ata cca aga gtc caa gtg gca caa att ctg ggt ccg ttg tcc atc aca	535
6014	Ile Pro Arg Val Gln Val Ala Gln Ile Leu Gly Pro Leu Ser Ile Thr	
6015	145 150 155 160	
6016	aac aag aca ttg att tat ata ttg gga ctg cag ctt ttc acc tct ggt	583
6017	Asn Lys Thr Leu Ile Tyr Ile Leu Gly Leu Gln Leu Phe Thr Ser Gly	
6018	165 170 175	
6019	tcc tac atc tgg att gta gcc ata agt gga ctt atg tcc ggt ctg tgc	631
6020	Ser Tyr Ile Trp Ile Val Ala Ile Ser Gly Leu Met Ser Gly Leu Cys	
6021	180 185 190	
6022	tac gac agc aaa atg ttc cag gtg cat cag gtg ctc tgc atc ccc agc	679
6023	Tyr Asp Ser Lys Met Phe Gln Val His Gln Val Leu Cys Ile Pro Ser	
6024	195 200 205	
6025	tgg atg gca aaa ttc ttt tct tgg aca ctt gaa ccc atc ttc tct tct	727
6026	Trp Met Ala Lys Phe Phe Ser Trp Thr Leu Glu Pro Ile Phe Ser Ser	
6027	210 215 220	
6028	tca gaa ccc acc agc gaa gcc aga att ggg atg gga gcc acg ctg gac	775
6029	Ser Glu Pro Thr Ser Glu Ala Arg Ile Gly Met Gly Ala Thr Leu Asp	
6030	225 230 235 240	
6031	atc cag aga cag cag aga atg gag ctg ctg gac cgg cag ctg atg ttc	823
6032	Ile Gln Arg Gln Gln Arg Met Glu Leu Leu Asp Arg Gln Leu Met Phe	
6033	245 250 255	
6034	tct cag ttt gca caa ggg agg cga cag aga cag cag cag gga gga atg	871
6035	Ser Gln Phe Ala Gln Gly Arg Arg Gln Arg Gln Gln Gly Gly Met	
6036	260 265 270	
6037	atc aat tgg aat cgt ctt ttt cct cct tta cgt cag cga caa aac gta	919
6038	Ile Asn Trp Asn Arg Leu Phe Pro Pro Leu Arg Gln Arg Gln Asn Val	
6039	275 280 285	
6040	aac tat cag ggc ggt cgg cag tct gag cca gca gcg ccc cct cta gaa	967
6041	Asn Tyr Gln Gly Gly Arg Gln Ser Glu Pro Ala Ala Pro Pro Leu Glu	
6042	290 295 300	
6043	gtt tct gag gaa cag gtc gcc cgg ctc atg gag atg gga ttt tcc aga	1015
6044	Val Ser Glu Glu Gln Val Ala Arg Leu Met Glu Met Gly Phe Ser Arg	
6045	305 310 315 320	
6046	ggt gat gct ttg gaa gcc ctg aga gct tca aac aat gac ctc aat gtc	1063
6047	Gly Asp Ala Leu Glu Ala Leu Arg Ala Ser Asn Asn Asp Leu Asn Val	
6048	325 330 335	
6049	gcc acc aac ttc ctg ctg cag cac tgatagtcac aggccaacac tgg	1110
6050	Ala Thr Asn Phe Leu Leu Gln His	
6051	340	
6052	gaccggaccg gcagccaggt gacagtgcgt ggteccacc atcagatcag cccggggacc	1170
6053	gagcatctct ggtgctgatg ttcttgtggg aagagggagg ttccaccgca cccctgacct	1230
6054	caaccgcaag actgttgccg ttttagtggt gagataagtt tgccattaca ttagcatgta	1290
6055	ttttctatct atatttttta ttgggcattt tccctaggtt ggagagtcag cactcgtttt	1350
6056	gaatgtgttt aaaatgcatt aaaatggaag atttctgcag gcagttgaat ggcactccag	1410
6057	atgggggaatt gctgtaaccc tcttactgta acatgtcctc tctgctgcg tgatggggag	1470
6058	agggtaaatgt tacttcacaa aggacatgtc agatccttct tcatggactt ttttagttac	1530
6059	tgttttttct ctcaaacttg ttttcgaatc tccctggagt gagggagaaa cagggagctg	1590
6060	aatcctcccc caagctgttc caggccagag gactctgcag taccttctcc tacatctagt	1650
6061	aacaaagaat ggtgataacc atgcactggt tcaaggttct ggagttctcc atgaaacttg	1710



## RAW SEQUENCE LISTING

DATE: 08/03/2001

PATENT APPLICATION: US/09/743,247

TIME: 16:29:05

Input Set : A:\ES.txt

Output Set: N:\CRF3\08032001\I743247.raw

6062	ggttaatttt	gctcagagta	tccggagtta	gccactagge	tgcgggtgaa	atgggatgga	1770
6063	gtagaacaac	agcaggcttc	ctggagccac	atgggctgac	tagggcactc	tgtggctggc	1830
6064	ctggcacggg	ctcagcccag	gaagaggaga	aacgatecct	tgcctgcccc	tcctgtggc	1890
6065	agggctaact	gcctggccct	cctggctcgc	agccagccag	ccccctggca	gcaggttctc	1950
6066	ctcagggtt	gggtcttcaa	cctgtggcga	caggaggcag	ggcagactgt	ggaggacagg	2010
6067	atgcagggtca	gggagagggg	aggcaggggt	ggaccgccat	gagcatgaaa	agaccggaag	2070
6068	caagttgact	cttgcaatgt	gcaactgtta	tgttctgcaa	aatgagcaac	gatgtatcaa	2130
6069	attgatgcaa	atttagatgt	tgatacttac	aataaagttt	ttaatgtgtt	tt	2182

## VERIFICATION SUMMARY

DATE: 08/03/2001

PATENT APPLICATION: US/09/743,247

TIME: 16:29:06

Input Set : A:\ES.txt

Output Set: N:\CRF3\08032001\I743247.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:523 M:283 W: Missing Blank Line separator, <220> field identifier  
L:550 M:254 E: No. of Bases conflict, LENGTH:Input:450 Counted:449 SEQ:21  
M:254 Repeated in SeqNo=21  
L:553 M:252 E: No. of Seq. differs, <211>LENGTH:Input:510 Found:509 SEQ:21  
L:559 M:283 W: Missing Blank Line separator, <220> field identifier  
L:601 M:283 W: Missing Blank Line separator, <220> field identifier  
L:674 M:283 W: Missing Blank Line separator, <220> field identifier  
L:740 M:283 W: Missing Blank Line separator, <220> field identifier  
L:778 M:283 W: Missing Blank Line separator, <220> field identifier  
L:827 M:283 W: Missing Blank Line separator, <220> field identifier  
L:888 M:112 C: (48) String data converted to lower case,  
L:910 M:283 W: Missing Blank Line separator, <220> field identifier  
L:946 M:283 W: Missing Blank Line separator, <220> field identifier  
L:986 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1618 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1690 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1749 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1853 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1889 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1942 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2025 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2085 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2121 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2169 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2780 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2858 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2911 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2997 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3057 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3099 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3189 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3227 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3373 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3412 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3565 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:93  
L:4104 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4192 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4270 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4285 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:113  
L:4285 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:113  
L:4332 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4414 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4529 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4541 M:254 E: No. of Bases conflict, LENGTH:Input:203 Counted:206 SEQ:116  
L:4617 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4668 M:283 W: Missing Blank Line separator, <220> field identifier

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/743,247

DATE: 08/03/2001

TIME: 16:29:06

Input Set : A:\ES.txt

Output Set: N:\CRF3\08032001\I743247.raw

L:4705 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4788 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5453 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5566 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5670 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5739 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5762 M:254 E: No. of Bases conflict, LENGTH:Input:340 Counted:330 SEQ:144  
L:5775 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5833 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5906 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5945 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5980 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5985 M:254 E: No. of Bases conflict, LENGTH:Input:58 Counted:55 SEQ:149  
L:6075 M:283 W: Missing Blank Line separator, <220> field identifier